DOCUMENT RESUME

ED 404 357 TM 026 148

AUTHOR Herman, William E.
TITLE An Analysis of Multiple-Choice Test Item Booklets.
PUB DATE Aug 96
NOTE 15p.; Paper presented at the Annual Meeting of the American Psychological Association (104th, Toronto, Ontario, Canada, August 9-13, 1996).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Higher Education; *Metacognition; *Multiple Choice Tests; *Responses; Testing; *Test Items; Test Results; *Undergraduate Students

ABSTRACT

Marks made by students on test item booklets were analyzed as a clue to better understanding of the metacognitive strategies employed during the completion of a 100-question multiple-choice final examination. Test item booklets of 56 undergraduates were scrutinized for the frequency of the following item markings; (1) no markings at all; (2) elaborations, which consisted of definitions, examples, etc.; (3) use of a "?" next to items; and (4) option elimination techniques involving marking out alternatives. All subjects marked at least one item on the examination, and overall 58% of the items were marked. Examination performance was found to be related to all marking variables as predicted, although only the lack of item marking relationship was found to be statistically significant. The results suggest that such item markings may be an important artifact of the test taking experience that can lead to better understanding of examination strategies and the overall test taking experience. (Contains two tables and seven references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 404 357

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

An Analysis of Multiple-Choice Test Item Booklets

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
WILLIAM E. HERMAN
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

William E. Herman
Associate Professor
Department of Psychology
State University of New York
College at Potsdam
Potsdam, New York 13676-2294
Office Phone: (315) 267-2610
FAX: 315-267-2170

Poster Session presented at the 104th Annual Convention of the American Psychological Association, Toronto, Ontario, Canada, August 9-13, 1996.

Running head: Test Booklet Analysis

TM026148

BEST COPY AVAILABLE

Abstract

Students' marks made on test item booklets were analyzed as a clue to better understanding the metacognitive strategies employed during the completion of a 100 question multiple-choice final examination. Some 56 undergraduate students' test item booklets were scrutinized for the frequency of the following item markings: no markings at all, elaborations (definitions, examples, etc.), use of a "?" next to items, and option elimination techniques (marking out alternatives). All subjects marked at least one item on the exam and overall 58% of the items were marked. Exam performance was found to be related to all marking variables as predicted, although only the lack of item marking relationship was found to be statistically significant. The results suggest that such item markings may be an important artifact of the test taking experience which can lead to better understanding of exam taking strategies and the overall test taking experience.

An Analysis of Multiple-Choice Test Item Booklets

Many researchers and theoreticians would be very interested to learn more about the metacognitive strategies used by students during traditional paper-and-pencil classroom testing experiences. Unfortunately, much of what transpires in the private minds of students during exam taking is poorly understood. The insights gained from such knowledge could help students adopt more effective test-taking strategies and reduce the influence of non-effective and anxiety provoking reactions during the test-taking experience.

Why should researchers examine the written comments of students during test taking? The markings on test item booklets recorded by students during an exam produce one of the few permanent artifacts of the test-taking experience that can easily be analyzed by researchers. While it is possible that stray marks might be made which are meaningless, it is more likely than not that the sense of urgency created by the test taking circumstance makes the majority of these markings meaningful, trustworthy, and worth analysis. It could be argued that even what a student believes is random doodling might be an unconscious outlet for a deeper form of mental work.

Only a few researchers have seized the opportunity to explore the written comments of students during exam taking in order to better understand the metacognitive processes employed by students during the actual taking of examinations. The

current study attempted to assess past efforts in this arena and extend the knowledge of markings made by students during exams based upon what is currently known in the field of cognitive psychology.

McKeachie, Pollie, and Speisman (1955) designed an experimental test-taking situation where half of the students were given answer sheets with spaces for comments and the other half were provided with traditional answer sheets. Although students in these groups did not express different feelings related to the test, the students who were allowed to write on their answer sheets scored higher on the test. This finding supported the view that allowing students to write comments improved performance because anxiety that had built up during the exam which interfered with memory recall and problem solving now had a release mechanism.

Calvin, McGuigan, and Sullivan (1957) also found that students who were allowed to write comments during an achievement test scored higher than other students in a control group and highly test anxious students made the greatest gains under these conditions. Smith and Rockett (1958) found an interaction with students' written comments and anxiety where written comments helped the performance of highly anxious students but hindered the performance of low-anxiety students. McKeachie (1963) concluded in a review of these and other studies that "reducing the stress of the examination by permitting students to write comments resulted in improved performance" (p. 1160).

In another later review which included the previously mentioned studies, McKeachie (1984) offered three possible hypotheses to explain these findings: (1) tension was reduced by a catharsis through written comments, (2) recall problems created a Zeigarnik effect where the unfinished task interfered with answering later questions, and (3) mental reorganization. McKeachie (1984) described mental reorganization as a circumstance where students are able to think more deeply about exam items and "writing the comment facilitates reorganizing their thoughts and provides cures for retrieval; thus, writing the comment may enable the student to choose the correct alternative" (p. 191).

McLain (1983) studied the test-taking strategies of students who varied in academic grades by using a "think aloud" technique. Outstanding students demonstrated the following behaviors at greater frequencies than other students: number of alternatives read, answers anticipated before reading the alternatives, incorrect alternatives critiqued or eliminated, and questions initially skipped.

Kim and Goetz (1993) explored very specific marking strategies for answering multiple-choice questions by analyzing the spontaneous marks made on students' test booklets. They found that nearly all students (95%) made one or more marks on the exam sheet. It was further found that students' test scores were related to option elimination marks and the use of marking strategies increased as items became more difficult on the exam.

The present investigation is an attempt to refine and expand the study of the test-taking strategies that students employ. The paucity of research in this area tells us that we have much to learn about the cognitive processes used during test taking, lack normative data on the frequencies of such student markings, and have just begun to identify the broad and specific categorical marking strategies used by students. Previous findings and theory would suggest that fewer items would be left unmarked and greater use of elaboration, use of question marks, and option elimination marks would be employed by high exam achievers. It is unclear whether the broad category of markings/no markings or more specific types of markings will be stronger predictors of exam performance.

Method

Subjects

A total of 56 undergraduate students in two educational psychology courses at a small, state-supported university in up-state New York served as subjects. This course was offered by the Department of Psychology and many students planned to become teachers, although many others hoped to enter another area of human services. Students in two different sections over a two-semester period participated in the study. The majority of the subjects were female students (75 % of the total sample) in the 18-23 year-old age range.

Materials

The final examination consisted of 100 multiple-choice

items. While a few exam items were designed to measure students' understanding of theory, concepts, and terminology, most exam items asked students to analyze a classroom/learning situation and demonstrate the application of knowledge. Kuder Richardson 21 estimates of reliability over several past final exam administrations for subjects averaged .83. Scores on the final exam ranged from 34 percent correct to 96 percent correct with a mean of 73% correct (median=75.00; standard deviation=14.18).

Procedure

Prior to taking the 100 item multiple-choice final exam, students were verbally encouraged to write any comments on the exam booklet containing the test questions as had been the procedure during four earlier classroom tests. Students were told throughout the course that some researchers believed that such markings on tests might help students concentrate on the questions and arrive at better answers. As a normal control feature of exam administration, students were asked to write their names on the test booklets before turning in their exams.

Test item booklets were inspected after the final exam for the following markings derived from previous research and detailed scans of the test item booklets: number of items with--no markings; elaboration (definitions, descriptions, or examples); question marks; and option elimination marks.

Results

The descriptive results of item marking patterns based upon different levels of exam performance are presented in Table 1.

All item marking categories appear to depict linear relationships. An inspection of the item marking patterns in each category for the high and low exam groups depicted mean values which support the research hypotheses.

All students marked at least one item and overall 58 percent of the items were marked by students in some manner. Frequency of marking exam items on the 100 item exam ranged from zero to 98. Some 14 students (25 % of the total sample) marked all 100 exam items in the test booklet. Since it was possible to mark 100 items on the exam (values could range between 0-100), the mean number of almost 42 items left unmarked on exams for the entire sample demonstrates considerable use of non-marking by test takers.

The substantial variance of most marking categories across all exam performance levels indicates that several distinct strategies were employed by students at different performance levels. The use of a marking strategy did not guarantee success on the exam and some successful students did not rely upon a consistent marking of the test booklet to achieve success.

Since the exam marking categories appeared to represent linear relationships among variables, correlational analyses were used to explore test marking strategy and exam performance. The intercorrelations between exam scores and the frequency of using the four test marking categories are depicted in Table 2. A statistically significant ($p<.01$) relationship was only found for the broad the marking category of number of items with no

markings and exam performance. Three additional intercorrelations between item marking strategies were found to be statistically significant.

Discussion

Every student marked at least one item in the test booklet. Although 58 percent of the exam items in the booklets received some sort of marking, a surprising number of items in the test booklets received no marks at all (42 %). When no markings were made it is apparent that students used a mental (non-notational) strategy to answer these questions. Differential strategies which lead to academic success need to be better understood.

These findings are generally supported by earlier research findings where "nearly every student (95 percent) marked at least one question" (Kim & Goetz, 1993, p. 215). Unfortunately, Kim and Goetz (1983) examined specific marking categories and did not compare the broad markings/no-markings on items with exam performance. The current study suggests that the frequency of using any marking strategy may be the superior predictor variable instead of specific marking strategies, since this broad marking category explained more variance in exam performance than specific strategies examined in the present study and the Kim and Goetz study.

The degree of intercorrelation found between option elimination, elaboration, use of no markings of items, and use of "?" can be understood in terms of the no marking/marking of items variable being inclusive of these other more specific forms of

markings. The marking/no marking of items variable is a broader category which is a better predictor of exam performance than the specific marking strategies of option elimination, use of the "?", and elaboration.

This broad variable of whether students provided any marks for an item would seem to encompass a wide range of cognitive strategies that may be in use during the test taking experience. The specific cognitive strategies which lead to any form of marking the test item booklets would seem to include the following situations: attempts to recall information, sorting out the meaning of such information, elimination of certain items from consideration, thinking more deeply about an item, questioning the accuracy of a response, confusions about the item, returning to an item at a later time in the exam, and identifying an option as the answer. The fact that students use different written symbols on the test item booklets for codes to these processes further complicates the uncovering of a clear picture of the process underlying test taking by using this methodology. For example, during the data analysis it appeared very clear that while some students employed a slash to signal option elimination, other students had employed the same slash as an indication of answer selection.

The understanding of test-taking strategy through the artifact of markings (or lack of markings) in test-item booklets is still in an exploratory stage. Since correlational studies cannot determine cause and effect relationships, caution is urged

in the interpretation of such relationships in such exploratory studies. Variables that can be identified which are related to success and failure can lead the way toward uncovering some of the mysteries of the test taking experience.

The current level of knowledge regarding test taking certainly supports encouraging students to mark their test booklets as tools to recall information, think more deeply about material, and select answers. This should only be a suggestion rather than a mandatory expectation, since some successful students seem to perform at high levels without depending upon such marking strategies. Test creators and examiners are also advised to allow students to use this strategy of writing on test-item booklets as a test-taking work pad. Only further research on this topic will clarify the many issues raised by this line of research such as: What is the normative data on broad and specific test item markings? What are the specific metacognitive and test marking strategies which lead to the broad category of whether an item is marked or unmarked?

References

Calvin, A. D., McGuigan, F. J., & Sullivan, M. W. (1957). A further investigation of the relationship between anxiety and classroom examination performance. Journal of Educational Psychology, 48, 240-244.

Kim, Y. H., & Goetz, E. T. (1993). Strategic processing of test questions: The test marking responses of college students. Learning and Individual Differences, 5(3), 211-218.

McClain, L. (1983). Behavior during examinations: A comparison of A, C, and F students. Teaching of Psychology, 10, 69-71.

McKeachie, W. J. (1963). Research on teaching at the college and university level. In N. L. Gage (Ed.), Handbook of research on teaching (pp. 1118-1172). Chicago: RandMcNally.

McKeachie, W. J. (1984). Does anxiety disrupt information processing or does poor information processing lead to anxiety? International Review of Applied Psychology, 33, 187-203.

McKeachie, W. J., Pollie, D., & Speisman, J. (1955). Relieving anxiety in classroom examinations. Journal of Abnormal and Social Psychology, 50, 93-98.

Smith, W. F., & Rockett, F. C. (1958). Test performance as a function of anxiety, instructor, and instructions. Journal of Educational Research, 52, 138-141.

Table 1

Descriptive Statistical Analysis of Item Marking Data Depicted in Terms of Exam Performance

	High Exam Group (n=10)	Middle Exam Group (n=36)	Low Exam Group (n=10)	Total Sample (n=56)
Final Exam	M=92.00 SD=2.98	M=74.51 SD=7.06	M=50.90 SD=7.70	M=73.14 SD=14.18
Broad Response Category				
No Markings	M=16.5 SD=31.55	M=41.77 SD=34.41	M=65.10 SD=33.22	M=41.70 SD=35.98
Specific Response Categories				
Elaboration	M=2.2 SD=2.97	M=1.71 SD=5.58	M=1.1 SD=2.42	M=1.66 SD=4.67
Use of "?"	M=3.2 SD=5.98	M=2.34 SD=5.58	M=0.3 SD=0.68	M=2.11 SD=5.10
Option Elimination	M=59.4 SD=61.72	M=38.31 SD=54.10	M=23.4 SD=20.64	M=38.98 SD=51.32

Table 2

Intercorrelations Between Item Marking Categories and Exam Performance.

	2	3	4	5
1. Exam Score	-.37 *	.04	.24	.20
2. No Markings	-	.28	-.41 *	-.40 *
3. Elaboration		-	.02	.45 *
4. Use of "?"			-	.02
5. Option Elimination				-

* p<.01